DOCUMENT RESUME

ED 115 765 CE 005 586

AUTHOR Hanushek, Eric A.
TITLE Regional Differences in the Structure of Earnings. Program on Regional and Urban Economics, Discussion Paper No. 66.
INSTITUTION Harvard Univ., Cambridge, Mass. Program on Regional and Urban Economics.
SPONS AGENCY Economic Development Administration (DOC), Washington, D.C. Office of Economic Research.
PUB DATE Oct 71
NOTE 64p.

EDRS PRICE MF-$0.76 HC-$3.32 Plus Postage
DESCRIPTORS *Demography; Educational Background; Geographic Distribution; *Geographic Regions; *Income; Labor Market; Mathematical Models; Racial Factors; *Surveys; Tables (Data); *Veterans

ABSTRACT

Slightly over 180,000 draftees were surveyed 10 months after leaving the military to gain data regarding training, employment, occupation and wages if working, marital status, education, Armed Forces Qualification Test (AFQT) score, age, race, military occupation, and home. Differences in earnings functions among smaller, homogeneous labor markets are examined in terms of urban and regional location and black and white individuals. Aggregate statistics about the structure of earnings, characteristics of earnings functions, regional variations in earnings, and black-white differentials are analyzed; 10 tables supplement the discussion. Conclusions state the estimated effect of schooling and ability differences is low, with an additional year of schooling worth about 5 percent higher income and one decile move in the AFQT score matched by 1 percent increase in earnings for whites, lower for blacks. The estimation of separate earnings functions for labor markets across the country is significant, with the choice of region for employment often equivalent to the marginal earnings of several years of schooling. Racial difference in earnings stems from differences in the earning functions for blacks and whites, not in schooling, ability, or experience levels. Sample and population characteristics are discussed in the appendix. The nonrepresentative character of the sample is emphasized. (LH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED115765

PROGRAM ON REGIONAL AND URBAN ECONOMICS

Discussion Paper No. 66

REGIONAL DIFFERENCES IN THE STRUCTURE OF EARNINGS

by Eric A. Hanushek

Harvard University

October 1971

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

CE 005 586

*This program is supported with funds from the Office of Economic Research of the Economic Development Administration in the Department of Commerce, Project Number OER-015-G-71-14. The authors take full responsibility for the views contained in the paper which are not necessarily those of the sponsoring organization.

REGIONAL DIFFERENCES IN THE STRUCTURE OF EARNINGS

by Eric A. Hanushek

There have been a number of attempts at estimating the relationship between schooling and earnings of individuals.[1/] The most common feature of these studies has been severe data limitations which have tended to dictate how the analysis could proceed. One of the most serious restrictions imposed by the data has been the assumption that earnings relationships are the same across the nation or, at least, across very sizable aggregations of states. This paper delves into the viability of such assumptions by looking at differences in earnings functions among smaller, more homogeneous labor markets.

The availability of a large sample of individuals which contains data about earnings, human capital characteristics and detailed geographic location allows estimation of separate earnings functions for different major metropolitan and "rural" remainder areas of the country. These estimates can be aggregated to yield estimates of

*This research was supported by funds from the Office of Economic Research of the Economic Development Administration in the Department of Commerce, Project Number OER-015-G-70-4 and OER-015-G-71-14. Data and processing support was provided by the Office of the Assistant Secretary of Defense for Manpower and Reserve Affairs. Special thanks go to Eli S. Flyer of that office. This research benefitted greatly from early discussions with John F. Kain and Zvi Griliches and from comments by Finis Welch and David O'Neill on an earlier draft. Responsibility for the final product rests, of course, with the author.

1/ A survey of past analyses can be found in Griliches [5].

"national" returns to schooling and can further be used to describe an economic potential surface for different geographical regions.

Finally, these estimated models provide some insights into the differences in incomes by race. Earnings differentials by race can be depicted as a function of input differentials, structural differences in earnings functions and differences in the geographic distribution of individuals. This analysis allows measurement of the relative importance of each of these factors.

I. Data and Models

The measurement of educational levels and abilities of individuals has always been quite imprecise. Census data (as used, for example, by Becker[1] and Hanoch[7]) contain many built-in limitations such as no knowledge of quality differences in either individual ability or schooling.[1/] Representative samples which provide information on incomes along with more accurate measures of human capital embodied within an individual have not, however, been plentiful. This had led to the use of more or less nonrepresentative samples, but ones which provide some insights into the roles of education, ability and experience

1/ Some attempts have been made to introduce school quality information by either regional stratification or use of other data sources (cf. Weiss [16]). However, individual census data from the 1/1000 sample only allow breakdowns by large regions and, thus, only allow the crudest division by quality.

in determining incomes (e.g., Hansen, Weisbrod and Scanlon[8] and Griliches and Mason [6]). This study falls into the latter category.

All enlisted men leaving the military during Fiscal Year 1969 were surveyed ten months after departing the military. 1/ The survey provided information on training, employment, occupation and wages if working, and marital status which was merged with service information about education, Armed Forces Qualification Test (AFQT) score, age, race, military occupation, and home of record of the individual. A sample of slightly over 180,000 individuals was selected from this data bank by choosing individuals who: (1) had been in the Army two years or less; 2/ (2) had completed the survey with respect to income; and (3) had been working full time. 3/ The restriction on length of service was applied to reduce the complexities that might be introduced by differing military experiences and, thus, differing levels

1/ This must be qualified as individuals were surveyed if they had a reserve commitment upon leaving active duty. Almost everybody with less than six years active duty will be in this position.

2/ For practical purposes this implies that the sample is composed entirely of draft inductees. Enlistees agree to remain in the Army at least three years.

3/ From the sample of those meeting the first criteria and connected with the labor force, 92 percent were working full time, 2 percent were working part time, and 6 percent were unemployed but looking for work. A portion of the part-time employees are in school and, thus, are not strongly connected to the labor force. Other analyses are being conducted for the remaining categories with special emphasis upon unemployment.

of training. While the impact of Army training on civilian earnings is an interesting topic, no explicit information on length or depth of military schooling is available. There is more variance in Army training experiences with longer times in military. By developing a more homogeneous sample in terms of military experience, it is possible to concentrate upon the relationship between earnings and education, ability and labor market structure.

Because the military draft has never been considered to be unbiased in its selection of individuals, sampling distortions must be considered here. Also, there is the possibility of systematic nonresponse to the survey which would yield greater divergence from the population as a whole.[1/] Detailed comparisons of the sample and national population distributions of educational attainment, ability scores and mean earnings are presented in Appendix A. As would be expected from the original Selective Service choice of draftees, the variance of schooling and ability scores is less in the sample than in the population as a whole. The mean values of schooling completed (12 years), AFQT percentile (53.5) and earnings ($7,031 annually) do not appear unreasonable for this age group, however.

1/ For the entire Army sample (regardless of length of service), the survey response rate was 73.9 percent. It is known that this nonresponse tended to be systematic by race and amount of education.

Even allowing for some sampling biases, it is important to note that both ends of the educational and ability distributions do appear to be adequately represented. For the purposes of estimating relationships between earnings and characteristics of individuals, variation in the observed characteristics is more important than a perfectly representative sample of individuals. Unless any sampling biases are correlated with important but unmeasured characteristics of individuals within the sample, the sample provides a good data base for investigating the underlying earnings relationships.

Models of earnings by individuals tend to be very simple, and this work represents no departure from that precedence. The basic model of earnings is:

(1) Earnings = f(education, ability, experience).

Several different forms of this basic earnings model were examined in the course of the estimation. Earnings defined as both hourly wages and weekly wages and in both linear and logarithnic functional forms were considered. The presentation here concentrates upon models which describe the log of weekly income as a linear function of the human capital attributes of the individual. Weekly income was used since human capital attributes would also be expected to affect hours worked which is certainly an important consideration in calculating

returns to education and training.[1] The choice of log wages was arrived at from experimentation with different functional forms.

Operationally, the human capital attributes on the right hand side in Equation 1 are jointly measured by years of schooling, AFQT percentile, and age of the individual. Since, however, we wish to measure the work experiences of the individual and not his age *per se*, we must make some assumptions about his pre-Army experiences. The simplest assumption is that all time outside of school and the Army was spent gathering relevant work experience.[2]

Thus, our desired model is:

$$(2)\quad \log Y = a + bS + cAFQT + dE$$

where Y is weekly earnings, S is years of schooling, AFQT is the percentile score of the individual, and E is experience.

1/ One concern in using weekly earnings instead of hourly earnings is that aggregate economic conditions could contaminate the model. The observations were recorded over a twelve-month period of changing economic conditions. The unadjusted national unemployment rate changed from a low of 2.9 percent to a high of 4.7 percent. In order to allow for this possible factor, models were estimated which included, as one of the indepedent variables, the unadjusted national unemployment rate during the month in which the individual answered the survey. This was invariably insignificant according to traditional statistical tests. Other analyses of these data indicate that whether or not a person is employed depends upon aggregate conditions. These models indicate no sensitivity of earnings to aggregate conditions, given that the individual is employed. An alternative mode of analysis would be the development of structural models for both hourly earnings and hours worked. This seemed much more difficult and also would require more data - particularly in the hours worked models.

2/ The effect of making the age-experience transformation was pointed out to me in an earlier draft by Finis Welch.

From estimating the model

(3) $\log Y = a^* + b^*S + c^*AFQT + d^*AGE$

and assuming that [1/]

(4) $E = Age - S - 8$,

we find that the desired coefficients are given by:

(5) $b = b^* + d^*$,

(6) $d = d^*$, and

(7) $c = c^*$.

The analysis presented here discusses only coefficients from the experience specification of Equation 2 under the assumption about experience in Equation 4.

The relative richness of the sample does allow testing several interesting extensions of the standard human capital hypotheses. Since measures of income pertain to ten months or less of job experience, it is likely that many of them are in some training status. As has been developed previously (e.g., Mincer[13]), people undergoing training by the firm would be expected to receive lower wages. The extent of this can be tested, albeit crudely, since the survey recorded whether or not individuals were undergoing training at the time of survey.

1/ This assumption can be relaxed in an aggregate manner later by relying upon teenage unemployment rates for groups in the economy (particularly racial groups), but no specific individual information on past work experiences is available.

Information is available from service records to indicate the military occupation of the individual. These data, giving roughly a one-digit occupational breakdown, allow testing for differential transferability of various military skills.[1/]

At the same time, data are available on the civilian occupation of the individual. This knowledge, again at the one-digit level, provides the ability to analyze occupational differences which exist over and above human capital or training differences. Such differences could exist through differences in the monopoly position of either labor suppliers or demanders. For example, the control of labor supply by the building trade unions would be expected to raise the wages for those who enter the included trades; that is, raise wages over what would be expected for a given level of human capital.

A central part of this analysis concerns the homogeneity of labor markets throughout the country. For data reasons past analyses have made very strong homogeneity assumptions about labor markets; in particular, they have assumed the same returns to human

1/ The one-digit occupational groupings are: (0) infantry, gun crews, and seamanship specialists; (1) electronic equipment repairmen; (2) communications and intelligence specialists; (3) medical and dental specialists; (4) other technical and allied specialists; (5) administrative specialists and clerks; (6) electrical/mechanical equipment repairmen; (7) craftsmen; (8) service and supply handlers. The data file does not contain information on length of training within each occupational field. An assumption must be made that individuals within the same one-digit field receive equal amounts of training in order to use this information. This is probably not a bad assumption for first term draftees but becomes increasingly tenuous as individuals are in the military for longer periods of time.

capital over very large regions. The common treatment of geographic location within the country has been to use large region intercept dummy variables (cf. Griliches and Mason [6]) or to stratify on very large regions (cf. Hanoch[7]). These crude techniques always display large and [illegible]nt differences in earnings by regions. There is, howev[illegible] reason to believe that these methods go far enough in accounting for differences in labor markets across the country. For rural areas, Welch [17] presents evidence that there are significant differences by states. Simple consideration of the workings of labor markets would suggest that macro adjustments are not enough: information flows - key to labor market adjustments - appear inadequate within these macro regions; further, industrial structure and input compositions differ considerably by locality. It might be tempting to appeal to factor price equalization theorems with trade among regions (e. g., Samuelson [14]). However, these theorems apply only to equilibrium, and there is little reason to believe that equilibrium has been obtained or even approached in this situation. The large internal migration streams offer prima facie evidence that this is not the case.

The question of labor market homogeneity can be approached in detail here because of the size of the sample available. Since the residence of each individual in the sample is known, individuals can be divided into fairly precise regions. The criteria for forming regions were as follows: (1) except in the South, all SMSA's with over 200,000 people in 1960 were considered separate regions; (2) in the South, all SMSA's were considered regions; (3) remaining areas [not regions by (1) and (2)] were grouped by states into 24 rural regions. This division of the country yielded a total of 165 regions. Within this framework, some preliminary estimates of the importance of labor market conditions on earnings can be analyzed.

For each of the labor markets (defined above), separate earnings models were estimated for both blacks and whites.[1/] There remains considerably question about differences in incomes by race. Is it all accounted for by differences in human capital, or does a significant proportion of the mean differences relate to discrimination in the labor market? With different models for each region, it is possible to analyze differences in the estimated earnings functions by labor market and by race. These models can be subjected to the usual statistical tests for homogeneity. Further, this sample stratification

1/ This is subject to some regional sample size considerations delineated below.

allows a more detailed analysis of regional income differences than has been possible in the past.

Initial analyses of the data indicated that there was a real problem in observed variance of the inputs. A fairly large number of individuals is needed within a sample in order to obtain good parameters estimates; that is, small regions may contain too little variation in the observed characteristics of individuals. A rather arbitrary sample size cutoff of seventy-five observations was placed upon the individual regional samples. The result of this was the elimination of fifteen SMSA's from the white urban samples, 114 SMSA's from the black urban samples, and eight regions from the black rural samples. This was, however, an elimination of only two percent of the individuals in the sample.[1/] This leaves 24 rural white, 126 urban white, 16 rural black and 27 urban black samples from which separate regression models are estimated.

II. Aggregate Characteristics

The individual regional models of earnings are all displayed elsewhere[11]. This section presents some aggregate statistics about the structure of earnings. The remaining sections discuss the

1/ The effect of the sample size limitation was more pronounced in the case of blacks. For that group, the loss rate was 19 percent, as compared with one-half percent for whites. Further, data problems caused the elimination of San Francisco and Philadelphia from the black analysis.

implications of the models as they relate to individual earnings and regional labor markets. A final section discusses differences in incomes for blacks and whites.

A conclusion that is made abundantly clear by all analyses of individual incomes is that much remains to be explained. The crude human capital models fail to explain as much as half of the variation in incomes.[1/] The data set for this analysis contained, relative to most past samples, a much richer view of the individuals in the sample. Even so, the results are far from overwhelming.

For descriptive purposes, the set of regional analyses is divided by urban and regional location and by black and white individuals. An overall summary of the performance of the models is displayed in Table 1. The variation in the log of individual weekly earnings is partitioned into the proportion of individual variation within regions and the variation between regions, or the variation in regional means. The proportion of the within region variance which is explained by the set of regional earnings regression models is displayed in Column 3 of the table.[2/] Finally, the sum of Column 3 times 2 and Column 1 yields a crude estimate of the aggregate explained variance

1/ This is also the case in all other modelling efforts for individual earnings; see, for example, [6] and [8].

2/ This is calculated as the total explained sum of squares over the total sum of squares for all of the regions. Thus, it is a weighted average of the R^2's in the regional models.

since Column 1 is the amount "explained" by stratifying the sample into regions, Column 3 is the amount explained in each of the regions, and Column 2 is amount of within region variance which could be explained by the regression models.

As indicated by Column 4, the regional models explain between an eighth and a quarter of the variance in incomes in each of the aggregation. This actually sets an outer bound on our knowledge of earnings relationships because about one half of the explanatory power lies in the regional division of the sample. The causes of such regional variations in mean income are not well understood -- although a later section of this paper does delve into part of the explanation.

The aggregate statistics of Table 1 also indicate some other features of the models and the sample. The extent of mean earnings differences by region appears from Column 1 to be considerably higher for blacks than for whites. On the other hand, the earnings models for whites tend to explain more of the within region variation in income. The difference in means for blacks could be particularly important since, if not due to input differences, it has strong implications for migration among regions. This will be analyzed in subsequent sections.

Finally, models from this sample of new entrants into the job market would be expected to possess lower explanatory power than ones

Table 1

ANALYSIS OF VARIANCE OF INDIVIDUAL EARNINGS[a/]

Grouping	# regions	# individuals	Proporation of Variance between 1	within 2	explained 3	4=1+3·2
All	193	180,330	.158	.842	.091	.235
Urban	153	72,882	.135	.865	.118	.237
Rural	40	107,448	.126	.874	.079	.195
White	150	168,069	.067	.933	.092	.152
Urban	126	65,599	.083	.917	.122	.195
Rural	24	102,470	.046	.954	.080	.122
Black	43	12,261	.195	.805	.068	.250
Urban	27	7,283	.129	.871	.074	.193
Rural	16	4,978	.158	.842	.060	.209

a/ Individual regional models for the log of income from which this table is derived are displayed in Hanushek [11].

from other conceivable samples taken later in individuals' work profiles. Since the survey information applies to a time ten months after separation from the Army, the earnings figure almost certainly contains a sizable transitory component. This would have the effect of increasing the unexplained error in the models. Thus, some of the advantages in descriptions of the individuals from using this sample are offset by the early survey date.[1/]

II. Characteristics of the Earnings Functions

The choice of regional divisions within this analysis was not the only one which could be made. In particular was it necessary to go to the fineness of regional definition used here, or were more aggregate regions of the type used in the past satisfactory? This question was looked at in some detail during the course of analysis, and the answer was clear. This detail is warranted. The appropriate covariance tests were applied to aggregate regions, and homogeneity

1/ The ten-month period does seem long enough to minimize one worrisome source of transitory earnings. Many individuals return to school after separation from the Army. The transitory component of earnings would be particularly high if the sample included many people in temporary jobs while waiting entrance into school. After ten months, one would not expect many still waiting to enter school. (Full-time students have been excluded from the sample so there is no contamination from temporary jobs held by students.)

within broad regions was consistently rejected at the one percent level.[1/] This was even the case when each of the micro regions of this study was allowed to have its own intercept.

When considering the point estimates of the regression parameters, it is useful to assess the precision of these estimates. Toward this end, Table 2 presents an abbreviated frequency distribution for the standard t-statistic of coefficients across the regional models. From this, according to traditional significance tests, there are very mixed levels of significance. The schooling coefficients are consistently well estimated. For whites, the estimated experience coefficients are also quite precise. A large part of the imprecision that does arise in the estimates appears to arise from lack of enough variation in the inputs. Larger samples--and ones with more observed independent variation in the exogenous variables -- consistently have more precise estimates. Had the sample size cutoff been set higher than 75 observations, the proportion of significant coefficients would rise dramatically. The trade-off of imprecision for larger

---

1/ The covariance, or Chow, tests used are described in Fisher [3]. The country was divided into seven aggregate regions. Separate covariance tests for urban and rural, white and black were performed for each of the seven regions. The aggregate regional groupings are displayed in Map 1 below. As an example of the F values, the core South had the lowest F-statistic for the seven rural white tests with $F(30, 13056) = 2.05$. (Intercepts were allowed to vary by each of the seven states in the aggregate test; if constrained to the same intercept, the same test yields $F(36, 13056) = 14.61$.)

Table 2

DISTRIBUTION OF t-STATISTICS

Coefficient	$t < 1.0$	$1.0 \leq t < 1.67$	$t \geq 1.67$
White Models			
Schooling	4	13	133
AFQT	51	30	69
Experience	8	7	135
Black Models			
Schooling	2	7	34
AFQT	24	11	8
Experience	16	9	18

number of regions seemed warranted.[1/]

The difference in earnings structure by labor market implies that it is not possible to present a single point estimate of the marginal importance of additional inputs. It is really necessary to think in terms of a distribution of returns dependent upon geographical location. In order to place the models in perspective, however, average parameters are presented for varying aggregations of the models (urban/rural, black/white).

Schooling. The majority of attention in earnings analysis has gone to the return to formal schooling. This is natural since such relationships indicate whether or not the optimal level of schooling is being purchased. The distribution of the individual schooling parameter estimates is displayed in Table 3. Since they do demonstrate such a large variance, subsequent discussions of average values must be tempered by this knowledge. The mean parameters presented in Table 3 are very much

1/ As an example of the effect of sample size, if the additional 16 black regions with between 75 and 125 observations were eliminated. The number of coefficients with t-statistics less than 1.0 would go from 2 to 1 for schooling, from 24 to 14 for AFQT score, and from 16 to 7 for age. Similarly, the t-statistics between 1.0 and 1.67 would go from 7 to 5, 11 to 8, and 9 to 5 for the respective variables. While increasing the size cutoff eliminates only 11 percent of the number of black observations, it eliminates 36 percent of the black regions. This reduction in number of areas led to the decision to retain the "medium" sized regions even though the coefficient estimates were not very precise.

Table 3

FREQUENCY DISTRIBUTION AND MEANS OF SCHOOLING COEFFICIENTS

	Value						Mean	
Grouping	<.020	.020-.034	.035-.049	.050-.064	.065-.079	.080	Raw	Weighted
All	6	29	54	61	28	15	.052	.049
Urban	5	23	42	43	25	15	.053	.052
Rural	1	6	12	18	3	0	.047	.046
White	4	23	39	48	22	14	.053	.049
Urban	4	20	31	36	21	14	.055	.053
Rural	0	3	8	12	1	0	.049	.046
Black	2	6	15	13	6	1	.049	.047
Urban	1	3	11	7	4	1	.051	.050
Rural	1	3	4	6	2	0	.045	.043

a function of the aggregation scheme. If the observed distribution diverges significantly from the geographic distribution of the population, a new weighting scheme is called for. Means for different aggregations are presented both in raw form and weighted by the number of observations in the regions. Since the models relate the log of earnings to years of schooling, the coefficients (times 100) are interpreted as the percentage increase in earnings that would result from a one-year increase in the quantity of schooling. 1/

Some caution is advised in the strict interpretation of these results. The appropriate rate of return for judging investments in schooling is based upon a causal relationship between additional schooling and additional earnings. If more able or more motivated individuals tend to continue longer in schooling and these abilities or motivations lead to increased earnings, estimates of earnings as a function of only schooling and experience would overstate the return to schooling. This was the reason for including the AFQT percentile in the models. The AFQT percentile, however, does include some achievement that resulted from schooling and does not accurately portray motivational factors. 2/ The first caveat indicates the

1/ It is difficult to compare these estimates directly with those obtained elsewhere because of different model specification. For example, Griliches and Mason[6] present age adjusted, rather than experience adjusted results. The estimated age adjusted results here are less than in Griliches and Mason. However, the higher estimated returns to experience bring the experience adjusted rates closer to their implicit estimates.

2/ For a description of these tests, see Karpinos[12].

coefficients in Table 3 tend to be underestimated while the second indicates that they may be overestimated. Some feel for the extent of the first bias can be gained from estimating models without the AFQT variable. Table 4 shows the weighted mean schooling coefficients in models with and without AFQT. For the total sample, the inclusion of the AFQT variable yields a fourteen percent reduction in the mean schooling relationship. This represents the maximum effect which could be attributed to an overall school quality effect; that is, a component of the AFQT score that is linearly related in the region to the level of schooling.[1/] What fraction of this is due to a causal relationship from ability to level of schooling cannot be determined. Any upward biases in the schooling coefficient from motivations which are unrelated to schooling also cannot be determined within this sample. Thus, while we would like to interpret the schooling coefficient as an "ability free" measure of schooling returns, there are some ambiguities still present.

With this interpretive qualification in mind, it is possible to illustrate the economics of the schooling decision which is implied. For the total sample mean schooling coefficient of .049 at the mean income level an additional year of schooling is worth $345 in annual earnings.[2/]

1/ The relationship between this ratio and heterogeneity of school quality will be discussed in the next section.

2/ This is calculated as the weighted mean coefficient times mean annual earnings in Table A-3 for the grouping assuming the person is paid for 52 weeks.

Table 4

Weighted Mean Schooling Coefficients with and without Variable

Grouping	with AFQT (1)	without AFQT (2)	(1) / (2)
Total	.049	.057	.86
Urban	.052	.059	.88
Rural	.046	.055	.83
White	.049	.057	.86
Urban	.053	.060	.89
Rural	.046	.056	.82
Black	.047	.049	.96
Urban	.050	.052	.97
Rural	.043	.045	.96

The corresponding figures for whites and blacks are $348 and $291 respectively. Calculating the advisability of another year of schooling of course depends upon how these starting salary differentials are capitalized into lifetime earnings. For illustrative purposes, an earnings-experience path similar to past observations was assumed; namely, earnings grow at the rate given by the estimated experience coefficient for 25 years, remain constant for 10 years, and decline at a 2.5 percent rate for seven years.[1/] The mean present value of an additional year of schooling with a five and a ten percent discount rate for this profile is displayed in Table 5.

Even at a five percent discount rate, a decision to obtain an additional year of schooling above the mean by blacks appears question-

---

1/ The profiles assumed in these calculations imply that people work on average to age 65. The present value calculations are not very sensitive to the exact break points in the profile but are sensitive to the beginning earnings differentials and to large differences in the experience coefficients.

This profile roughly follows the schooling-age profiles displayed by Hanoch[7]. An interesting aspect of them, however, is how closely they seem to follow a stationary experience relationship. The earnings for increasing amounts of schooling tend to peak at older ages; also, people with more schooling tend to work longer in life. (This latter factor is indicated by an increasing ratio of expected earnings at age 77 to age 67 by years of schooling.) If age profiles are adjusted for differing amounts of experience due to differing amounts of schooling, the peaks and lengths of the tails appear quite similar. This can be interpreted as further evidence that the experience specification is more appropriate than the age specification.

Table 5

Mean Present Value of an Additional Year of Schooling[a/]

Grouping	Discount Rate 5%	Discount Rate 10%
Total	8,764	5,021
Urban	9,349	5,356
Rural	7,947	4,530
White	8,979	5,125
Urban	10,275	5,844
Rural	8,104	4,644
Black	5,982	3,568
Urban	6,641	3,960
Rural	5,086	3,025

a/ All calculations are based on the weighted mean schooling coefficients in Table 3 and the weighted mean experience coefficients in Table 7. The assumed earnings profiles are described in the text.

able.[1/] The mean earnings in rural areas (which would be foregone by continuing schooling) is greater than the present value of an additional year of schooling. The present value of a year's schooling exceeds mean earnings for urban blacks by only $183, or less than one year of college tuition. For whites, an additional year of schooling appears reasonably with a five percent discount but dubious with a ten percent discount. (Mean earnings for each grouping can be found in Appendix A.) Again, since there are large regional differences in the estimated returns to schooling, these overall observations must be tempered somewhat.

The consistency of the estimated relationships at the extremes of the educational distribution was tested through the introduction of intercept dummy variables for individuals with a college education or more and individuals with less than a high school education. Neither of these variables proved to be significantly different from zero.

1/ Since these calculations use the sample means, the changes in education refer to changes around the mean. For both blacks and whites the mean levels of schooling are quite close to 12 years; thus, the calculations can be thought of in terms of the decision to go to college.

The generalizations about the advisability of another year of schooling do not apply precisely to the individuals in this sample. By virtue of being in the military, the GI Bill payments for schooling lessen considerably the opportunity cost of continuing in school.

Ability. Ability differences of individuals were measured by Armed Forces Qualification Test percentile scores. The coefficient estimates for this variable are consistently less precise than those for the other variable in the model. (See Table 2) This imprecision could arise from a number of sources. First, cognitive ability and achievement could have little or no impact on the earnings.[1/] Alternatively, this could be a very poor measure of the ability quantity which is important; the test could be unreliable (i.e., a large sampling error of the test) or the test could be invalid (i.e., it doesn't measure what it purports to measure).[2/] Within these data, it is not possible to distinguish adequately among the competing explanations.

Table 6 displays the frequency distribution and means for the estimated coefficients. There is less variance within groupings in these coefficients than in the schooling coefficients. The black coefficients are consistently much lower than the white coefficients. Estimates of ability parameters across labor markets would not be subject to aggregation problems of the same severity as schooling parameter estimations as long as samples are stratified by race.

1/ This hypothesis is best stated in Gintis [4].
2/ This hypothesis and an attempt to deal with it are contained in Griliches and Mason [6].

Table 6

Frequency Distribution & Means of Ability Coefficients

Grouping	Parameter Value						Means	
	<0	0-.001	.001-.002	.002-.003	.003-.004	.004-.005	Raw	Weighted
Total	27	64	72	24	4	2	.0009	.0012
Urban	26	55	52	14	4	2	.0009	.0009
Rural	1	9	20	10	0	0	.0009	.0013
White	17	47	59	22	4	1	.0011	.0012
Urban	17	45	45	14	4	1	.0011	.0010
Rural	0	2	14	8	0	0	.0013	.0013
Black	9	17	13	2	0	1	.0004	.0005
Urban	9	10	7	0	0	1	.0003	.0005
Rural	1	7	6	2	0	0	.0005	.0005

These estimates, which are very similar to those of Griliches and Mason [6], appear very small.[1/] They indicate that a decile change in position in the test score leads to only a one percent change in white earnings or a one half percent change in black earnings. This implies that one to two years of additional schooling is equivalent to moving the entire range of the ability scale in terms of the change in earnings.

1/ This is considerably different from the findings of Hansen, Weisbrod and Scanlon [8] for low achievers. They found that including an achievement measure reduced the schooling coefficient to insignificance.

The interpretation to be placed upon this coefficient by itself is clouded by the same concerns as the schooling coefficient. At this point, a more formal presentation of the problem will clarify the various aspects of it. First, let us decompose this measure of ability and achievement of an individual ($AFQT_i$) into a school component and a nonschool component.[1/] Further, as supported by other studies ([2], [9], and [10]), let the relationship between achievement and schooling depend upon the quality of the school attended by the individual. We can then represent $AFQT_i$ as:

(8) $AFQT_i = a_o + a_1\widehat{AFQT}_i + a_{2j}S_i + e_i$

where $\widehat{AFQT}_i$ is the nonschool component, a_{2j} is the quality coefficient for the jth school and e_i is a stochastic component. If we let $\bar{a}_2$ be the mean school quality in the region, we can rewrite (8) as:

(9) $AFQT_i = a_o + a_1\widehat{AFQT}_i + \bar{a}_2S_i + (a_{2j} - \bar{a}_2)S_i + e_i$.

If the AFQT coefficient in the earnings model is B, we would like to add Ba_{2j} to the estimated returns to schooling. In the present model formulation this is all attributed to the AFQT variable. Yet, in the models displayed in Table 4 estimated without AFQT, we likewise do not attribute the right amount to schooling. Instead, if we assume that

1/ For this development, we make a rather strong assumption that there is no interaction between these components.

length of schooling is independent of school quality, the amount attributed to schooling in the nonAFQT models is $(Ba_1 r + B\bar{a}_2)$ where r is the correlation between $\widehat{AFQT}_i$ and S_i. Further, variation due to differences in regional school quality is ignored when AFQT is excluded from the model. The observation that AFQT includes some schooling makes the size of the estimated ability component of earnings appear even less reasonable.

Tests were also made for the consistency of the ability relationship across the entire range. Both continuous AFQT (described above) and a series of dummy variables for different percentile intervals were analyzed. The continuous variable performed better than the discontinuous variables. Further, when intercept dummy variables for the top and bottom ranges of the test were added to the continous variable, they were not significantly different from zero.

Experience. The measure of experience of the individual has a consistently strong effect on earnings, especially earnings of whites. Since the length of time on the present job is roughly the same for everybody by virtue of being surveyed at the same time after separation from the Army, this coefficient can be interpreted almost entirely as a work experience coefficient.

The frequency distribution and means of these coefficients are displayed in Table 7. It is interesting that the white experience coefficient is almost double that for blacks. Part of this could arise from age being an imperfect measure of work experience. Since the unemployment rate for black teenagers is considerably higher than that for white teenagers (historically almost double), the same chronological age for a white and black is not associated with the same average work experience level.

Making a gross adjustment for racial differences in the teenage unemployment rate, however, is not enough to bring equality in the mean experience coefficients. We can make an alternative experience transformation of age to allow for these differences as follows:

(10) $E = p_i (AGE - S - 8)$

where p_i is the teenage employment rate for blacks or whites. This implies that the experience coefficient is not given by Equation 6 but instead by[1/]:

(11) $d = d^*/p_i$

Assuming an employment rate of .85 for whites and .70 for blacks

1/ Note that the calculations of the schooling and AFQT coefficients are invariant to such assumptions about employment rates.

Table 7

Frequency Distribution and Means of Experience Coefficients

Grouping	Value					Mean	
	<0	0-.020	.021-.040	.041-.060	.061-.080	Raw	Weighted
Total	7	59	88	35	4	.026	.028
Urban	7	45	64	33	4	.027	.028
Rural	0	14	24	2	0	.022	.027
White	5	27	79	35	4	.030	.029
Urban	5	24	60	33	4	.031	.030
Rural	0	3	19	2	0	.027	.028
Black	2	32	9	0	0	.013	.013
Urban	2	21	4	0	0	.012	.013
Rural	0	11	5	0	0	.015	.014

(which roughly corresponds to conditions in 1964-66), the employment adjusted mean experience coefficient for whites would be .034 compared to .015 for blacks. This factor, therefore, does not seem to be the explanation of the differences in experience records.

The remaining explanation for the different rewards to experience by race lies in job discrimination. In particular, if blacks are consistently placed in jobs with less room for future advancement than whites with similar educational levels, the returns to additional years of experience will be less. The exact magnitude of such occupational discrimination is difficult to assess, however, if one believes that there has been social progress toward eliminating discrimination through the 1960's. In such an event the more experienced people in the sample would have been in the labor market earlier in the sixties when discrimination was more intense. This would have the effect of biasing the experience coefficient for blacks downward. Even so, such biases probably do not account for all, or even a majority of the black-white differences since time on the current job is roughly the same for everybody in the sample, and it is doubtful that blacks consistently returned to that same job path that they had left over two years before.

Occupation. Along with the human capital information about the individual, there is information available about his civilian occupation. Through introducing intercept dummy variables for the one-digit level, two occupational classifications stood out as significantly affecting earnings; that is, having an independent influence of earnings after allowing for differences in schooling, AFQT percentile, and experience. These were agricultural jobs and structural jobs.[1/] Each of these was included in a region's model if the coefficient was significantly different from zero at the five percent level; otherwise, it was excluded.

The effect and explanation of these factors are quite different. Within 23 of the 24 rural white regions and 9 of the 16 rural black regions (comprising 99.1 and 52.8 percent of the total individuals), a significantly negative relationship between earnings and agricultural occupations was estimated.[2/] The weighted mean coefficients across all regions was -.216 for whites and -.145 for blacks. If, however, the

---

1/ Structural work includes most of the construction trades. Major subcategories are: metal fabricating; welding and flame cutting; electrical installation, assembly and repair; painting, plastering and cementing; excavating, grading and paving; construction work, n.e.c.; structural work, n.e.c.

2/ The term rural, it should be remembered, has a special meaning in the context of the regions for this analysis. Rural refers to all land area left after removing the included SMSA's. In all but the South, this rural includes anybody not in an SMSA of 200,000 people or more. The percentage of rural individuals in agriculture is fairly low -- 3.4 percent for whites, 1.6 percent for blacks.

regions which do not have an estimated agricultural effect are excluded, the weighted mean coefficient for blacks goes to -.275. In other words, ceteris paribus, someone in agriculture would be expected to have earnings one-fifth to one-quarter less than someone not in agriculture.

There are two likely explanations of this large earnings discount for agriculture. First, there is a problem of measuring income. If agricultural jobs provide considerably more income in kind than other jobs, the nominal earnings measures understate the real earnings in agriculture. Second, the agricultural sector could simply be a depressed wage market, although almost certainly not to the extent indicated by the estimated parameters.

In 56 urban white regions and 9 urban black regions (comprising 76.1 and 53.5 percent of the total individuals), significant positive effects were estimated for the structural trades. The weighted average of the coefficients in all urban regions was .065 for whites and .059 for blacks; within regions in which there were significant estimates, these means were .085 and .110. These estimates are very plausible given recent discussions of labor supply restrictions

by the building trade unions and sensational announcements of new contract agreements in the construction fields.

Training. The measurement of training is imperfect for an analysis of its effects on earnings. To be precise one would want to know the type of training -- specific or general, length of training, and whether firm sponsored or not. Instead, the survey data indicate simply whether or not the individual is receiving formal training. Moreover, with the short job experience it is possible to argue that virtually everybody is in a training program even though it is not specifically identified as such. Thus, estimation of the reduction in earnings for formal training involves considerable error.

In 22 of the 126 urban white regions and 2 of the 27 urban black regions the estimated cost of training was statistically significant at the 5 percent level.[1/] Within these regions the average estimated cost to the individual being trained was .092 and .086 for whites and blacks respectively. Comparing these coefficients with the experience coefficients yields the estimate that whites can recoup their costs of training in about three times the length of training while it takes blacks over six times the length of training. The estimates for training are not

1/ A higher percentage of individuals than regions were involved; 38 percent of the whites and 29 percent of the blacks resided in regions in which training proved to be significant.

as persuasive as the other estimates, however, as they apply to significantly fewer regions than those for other parameters and the measurement errors seem larger.

Other Factors. As mentioned earlier, several other hypotheses about earnings functions were tested. None of them proved to have a significant impact on earnings. Yet, they deserve mention if only to chart the ground that has been covered.

The military occupation of the individual -- measured at the one-digit level -- never displayed any independent impact on post-service earnings. This is surprising since, on the surface, one would expect a considerably difference in the transferability of such skills as combat infantry and electronic equipment repair. The most likely explanation of this is the low level of training which is provided inductees into the Army. It appears that only through original enlistment or re-enlistment after induction will the individual be given any advanced training with differential carryover to the civilian economy.[1/]

The marital status and family size of the individual were also considered but rejected as a significant factor in determining earnings.

1/ Since everybody served the same time in the Army, it is not possible to estimate the importance of Army experience relative to civilian experience. The estimates of Guiliches and Mason imply that a year in the Army is worth 95 to 97 percent as much as a year in civilian life.

While it is hard to explain the structural model which would lead to including these attributes, they have been included in other studies and found to be significant.

Finally, the military experience leads to considerable geographic migration. Between ten and twenty percent of the individuals in each region entered the military from a different region. (This again uses the 165 regions defined for this analysis.) A plausible hypothesis is that either more able individuals in an earnings sense or ones who have additional information about jobs are the ones who reside in new areas. This hypothesis was tested with an intercept dummy variable for those who returned to their region of prior residence. No systematic differences between migrants and nonmigrants were found.

IV. Regional Variations in Earnings

As shown in Table 1, almost 16 percent of the variance in earnings results from differences in the mean earnings among regions. Within the sample, blacks within the rural areas of the core South earn only 69 percent as much as blacks in the urban Great Lakes regions; blacks in the urban Northeast earn 88 percent of whites in the urban Northeast; and so on. Are these mean earnings differentials simply a reflection

of input differentials, or is the structure of earnings (the various model coefficients) the dominant factor?

The answer to this question comes from some manipulation of the expression for the variance of the means between regions. For the jth regions the estimated earnings function is:

(12) $E^j = X^j b^j + e^j$

where E^j is a vector of earnings in regions j, X^j is a matrix of inputs (education, AFQT, age, training, and occupation), b^j is a vector of estimated coefficients, and e^j is a vector of residuals. For the set of mean inputs in region j ($\bar{X}^j$), we find that

(13) $\bar{E}^j = \bar{X}^j b^j$

where $\bar{E}^j$ is the mean earnings level in region j.

Let $\bar{M}$ be the vector of national mean levels for the inputs into (12).[1/] Then, define

(14) $\hat{E}^j = \bar{M} b^j$

so that $\hat{E}^j$ would be the predicted mean earnings in region j with the national mean level of inputs. Then, letting NE equal the national mean earnings levels, n_j equal the number of observations in region j,

1/ Both $\bar{M}$ and NE refer to national means within the group being considered; for example, within black rural regions.

and T equal the total number of observations, we perform the following manipulations with the expression for the variance of the earnings associated with variance in the regional means:

$$(15)\quad \frac{\Sigma n_j (\bar{E}^j - NE)^2}{T} = \frac{\Sigma n_j \{ (\bar{E}^j - \hat{E}^j) + (\hat{E}^j - NE) \}^2}{T}$$

$$(16)\quad = \frac{\Sigma n_j (\bar{E}^j - \hat{E}^j)^2}{T} + \frac{\Sigma n_j (\hat{E}^j - NE)^2}{T} + \frac{2 \Sigma n_j (\bar{E}^j - \hat{E}^j)(\hat{E}^j - NE)}{T}$$

On the right hand side in (16), the first term compares in each region the predicted earnings for the mean regional inputs ($\bar{E}^j$) to the predicted earning with the mean national inputs ($\hat{E}^j$), using the earnings structure estimated for each region. Thus, this is the variance due to input differences, holding structure constant. The second term compares the estimated earnings in the jth region with the estimated earnings for the nation (NE), using the national mean inputs in each case. This is then the difference in earnings due to differences in earnings structure by region since the input levels are held constant. The final term is an interaction component reflecting whether individuals with above average input levels tend to locate in regions that pay above average (+) or vice versa (-). If we divide through (16) by the total variance in earnings [the left hand term in (15)], we are left with a

division of the variance of mean earnings into a proportion due to mean input differences (levels of education, etc.), a proportion due to structural differences in the earnings relationships (values of b^j), and a proportion due to the interaction of inputs and earnings structures.

The results of this decomposition of variance are shown in Table 8. The first column shows the proportion of the total variation in incomes which is explained by differences in the mean earnings levels among regions. (This is the same as shown in Table 1.) The remaining columns distribute this variance among different sources.

The implications of the table are clear. In no case do input differences account for more than six percent of the total variance in mean earnings.[1/] On the other hand, structural differences in the earnings relationships among regions account for over 80 percent of the variance in mean earnings. For the total rural sample and for the total black sample, the variation that would result from structural difference alone is greater than the total variance in mean earnings. The explanation for this apparent anomaly is simple: there is a negative interaction term (high earners located in low paying regions) which suppresses the variance from what would be observed if individuals were

1/ To be precise, this analysis refers to the variance in the log of earnings.

Table 8

Decomposition of Mean Earnings Variation

Grouping	Variation Between Regions	Decomposition		
		Input Differences	Structural Differences	Interaction
Total	.158	.028	.994	-.022
Urban	.136	.055	.806	.139
Rural	.126	.018	1.042	-.059
White	.067	.034	.837	.129
Urban	.083	.040	.839	.121
Rural	.046	.018	.811	.171
Black	.195	.013	1.016	-.029
Urban	.129	.018	.979	.039
Rural	.158	.006	.956	.038

located randomly. These results, remember, do not imply that the characteristics of an individual have no effect on his earning ability. They imply that regional differences in inputs within this sample explain little of the regional differences in earnings.

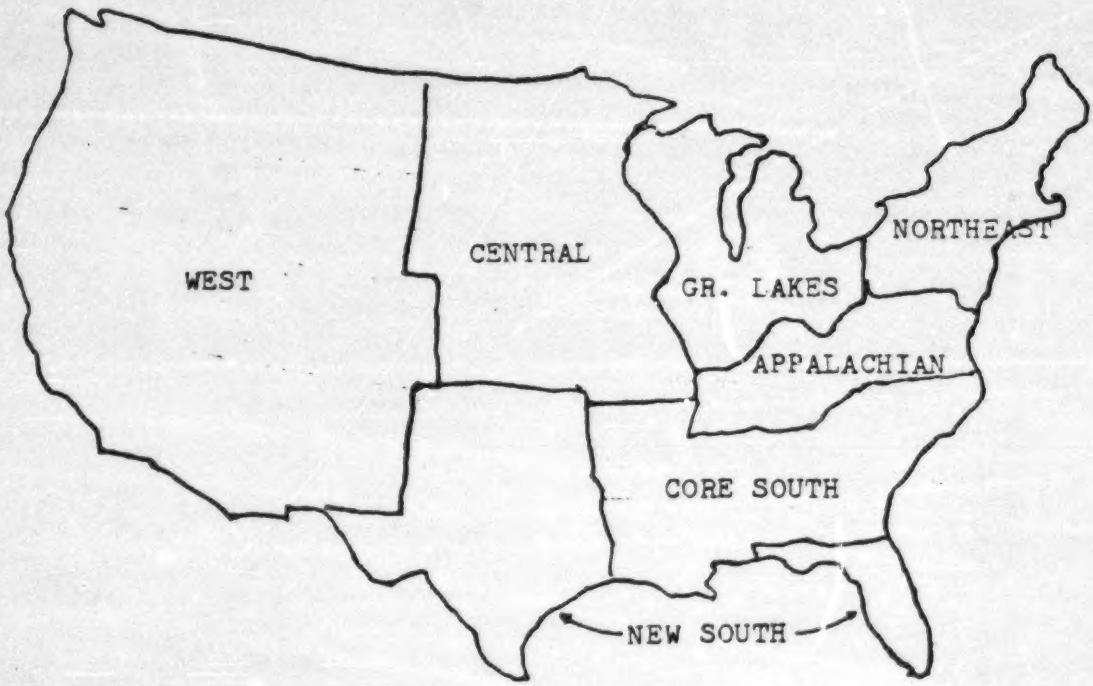
This suggests quite strongly that more effort should be devoted to analyzing the structure of labor markets than looking at the distributions of individuals and their characteristics in analyzing regional income patterns. Studies which account for variations in regional incomes by variations in aggregate education and experience levels overlook more basic, structural differences in the labor markets within each of the regions. Further, the differences in earnings structure among labor markets could provide important clues about internal migration patterns. When earnings potentials of individuals are estimated for macro regions or the country as a whole, it is difficult to see either how specific locational decisions are made within a broad region or how there is sizable migration in both directions. Differences in earnings structures within smaller regions could provide some of these answers.

Aggregate differences in the structural estimates can better be seen by grouping the estimated functions into macro geographic regions. Once this is done, the patterns of economic returns for education, ability

and experience become clearer. The grouping of states into seven macro regions is shown in Map 1. These regions, which do not correspond to the standard census divisions, were chosen in an attempt to group areas into more homogeneous economic regions. This was particularly true in the southeastern section where three regions were defined: Appalachian, Core South, and New South.

Table 9 displays mean weekly earnings and mean coefficients for education and age in each of the seven regions by race and place of residence. The distribution of ability coefficients does not show much variance and has, thus, been omitted from the table. (Remember, however, that there are significant racial disparities in them.) The observed mean earnings follow the pattern expected for the whole country: rural earnings are less than urban; black earnings are less than white; and southern area earnings are less than others.

The estimated returns to education follow a consistent pattern. The earnings functions are slightly steeper in the urban areas of a given region than in the rural remainder regions. Also, the returns to education in the three southeastern regions tend to be higher than elsewhere. This is very interesting especially when the large schooling quality differences between the South and elsewhere are recognized (see [2]).



MAP 1
MACRO REGIONS OF THE UNITED STATES

Table 9

Weighted Mean Structural Characteristics for Macro Geographic Regions

Group and Parameter	REGION						
	West	Central	New South	Core South	Appal.	Northeast	Gr. Lakes
Urban White							
Earnings[a/]	$139	$138	$124	$121	$126	$132	$146
Education	.051	.053	.057	.076	.058	.057	.051
Experience	.034	.026	.030	.044	.032	.029	.034
Rural White							
Earnings[a/]	130	122	117	112	114	128	136
Education	.031	.051	.057	.052	.056	.044	.008
Experience	.023	.029	.027	.026	.029	.017	.003
Urban Black							
Earnings[a/]	132	131	112	106	115	116	137
Education	.039	.060	.047	.051	.056	.052	.044
Experience	.014	.007	.006	.008	.019	.014	.013
Rural Black							
Earnings[a/]	--	--	101	95	102	124	134
Education	--	--	.049	.043	.053	.038	.039
Experience	--	--	.022	.011	.014	.022	.012

a/ Geometric mean of weekly earnings within the region.

Finally, in ten out of 12 comparisons, the marginal returns for an extra year of education are higher for whites than for blacks.

It is tempting to explain the regional differences in educational returns by differential demands for skilled labor (as between urban and rural) and differential supplies of educated labor (north and west versus south). However, the complexity of such explanations requires considerably more analysis than is feasible here. To do this correctly would require developing labor market models for the micro regions used in these estimates.

The experience parameters are not as consistent as the education parameters. The returns in experience tend to be slightly higher for urban areas than for rural areas. They also tend to be higher for whites than blacks. There are more exceptions to these observations than the ones about education, however.

IV. Black-White Differentials

There continues to be considerable interest in the relative incomes of blacks and whites. There have been numerous discussions of earnings differentials, e.g., Thurow [15]. In 1960, median black income for males was 52.6 percent of median white income.[1/] Within urban areas this figure was 58.3 percent; within rural areas it was 33.7 percent. This picture has not changed much since 1960. Given this, there is considerabl concern about the causes of these differences.

The overall picture from this sample does not look as bleak as that from national averages. In the aggregate, black earnings are 87.2 percent of white earnings while the figures for urban and rural are 87.3 percent and 82.0 percent, respectively. Part of this relative improvement in the sample is due to a slightly more favorable geographic distribution of sample blacks as opposed to the 1960 geographic distribution. However, the more important factor seems to be the relative closeness of input levels within the sample as opposed to the entire 1960 population. For example, the median years of schooling completed by male blacks over 25 years old in 1960 was 7.9 compared with 10.3 for whites. Within the sample the mean years of schooling completed is 11.89 for blacks as compared with 12.02 for whites. The

1/ U.S. Census of the Population, 1960. General Social and Economic Characteristics: United States Summary, p. 228-229. All 1960 population figures come from this source.

age and earnings range is also severely restricted. If discrimination is more severe in higher incomes which aren't adequately represented here, the sample differential will be reduced from the population figure. Such greater discrimination is indicated in the sample by lower experience coefficients for blacks. A final reason for the more favorable ratios within the sample is that only employed individuals are considered. Since blacks tend to be unemployed at a higher rate than whites, this ratio overstates the total income picture between the races.

Neveretheless, it is interesting to look at the differentials which do exist within the sample and attempt to identify the causes of these differentials. The previous discussion of the earnings models already has indicated that the structure of the models tend to differ systematically. by race. Blacks appear to receive slightly lower returns for marginal years of schooling and considerably lower returns for additional experience. This holds both in the aggregate and in the macro regions of the country. It is therefore valuable to ascertain what the structural differences in earnings relationships implies for aggregate black earnings.

It is possible to predict black earnings under different conditions. Since both earnings structures and input levels differ between blacks and

whites, separate estimates can be made of the effects of these differences. Within the 27 urban regions and 16 rural regions in which both black and white earnings models were estimated, two predictions were made: (1) mean black earnings from the black earnings models but using the white input mean characteristics for each region; and (2) mean black earnings from the white earnings models using mean black input characteristics. 1/

These predictions for black earnings along with the actual mean earnings for blacks and whites are displayed in Table 10. Looking at the last two columns, one can see the effects of input differences and structural differences by race. For rural areas, if blacks had the same characteristics as whites in each region, the disparity in earnings would remain the same; however, if they could receive wages according to the white earnings structure (without changing any input characteristics), 56 percent of the racial gap would be eliminated. In the urban areas, the predicted black earnings using white input levels reduces the earnings disparity from .87 to .89. However, receiving the same reimbursement for

1/ The estimates here all relate to geometric mean weekly earnings. They are weighted by the sample distribution of the black population. Looking at the reverse situation of the decrease in white earnings associated with black mean inputs and black earnings structure but weighting by the white population distribution makes only very slight changes in the predictions.

Table 10

Actual and Predicted Mean Earnings

	Actual		Predicted Black Earnings			
	White	Black	Black Struct., White Means[a/]	White Struct., Black Means[b/]		
	(1)	(2)	(3)	(4)	(3)/(1)	(4)/(1)
Rural	$126	$103	$103	$116	.82	.92
Urban	138	120	123	133	.89	.96

a/ Predicted geometric mean of weekly earnings using the black earnings function for each region and the mean values of white inputs in that regions.

b/ Predicted geometric mean of weekly earnings using the white earnings function for each region and the mean values of white inputs in that region.

their input characteristics as whites increase the earnings ratio to .96, or 69 percent of the earnings differential.

The picture is clear. The largest cause of differences in earnings between blacks and whites is a difference in the rates of reimbursement for skills and abilities (as reflected by education, AFQT, age, training status and occupation). Although blacks have lower schooling levels, lower AFQT levels and lower levels of participation in the high paying construction industries, these factors do not account for much of the difference in earnings.

Summary and Conclusions

Models of individual earnings still need considerable work. Simple human capital models do not capture a large part of the earnings potential of individuals. Nevertheless, while they might miss many important attributes of individuals, the models developed with existing knowledge and data do provide insights into the role of schooling and abilities in determining income.

The value of education or other inputs cannot, however, be described by a single statistic. Instead they appear to be a function of the geographical area in which the individual lives. Considering major metropolitan areas as separate labor markets, one finds significant variation in the returns to human capital across labor markets. This implies that past analyses of the returns to schooling, ability and experience will be very dependent upon the geographic distribution of the individuals in the sample and, thus, upon the specific aggregation of relationships for different labor markets.

It is difficult to sort out precisely the independent effects of schooling, abilities and motivation on earnings. Nevertheless, the direct estimates of schooling for constant AFQT percentile and experience level indicate that an additional year of schooling is worth about five

percent higher income. This figure implies that a decision to enter college by the average person in this sample is questionable - particularly if he is black.

The estimated effect of ability differences is likewise very low. A decile move in AFQT score (which contains some school effects) is matched by only a one percent increase in earnings. This small effect, while found by others, still confronts a priori beliefs.

The estimation of separate earnings functions for labor markets across the country provides a picture of the economic potential surface facing each individual. The effect of structural differences in the earnings functions is dramatic: over 80 percent of the differences in mean earnings among labor markets is attributable to differences in earnings structure as opposed to differences in input means among regions. When magnitudes of mean earnings differences among regions are considered, it becomes evident that choice of region is very important and is equivalent in many cases to the marginal earnings of several years of schooling.

At the same time, differences in earnings by race appear to arise fundamentally from differences in the earnings functions for blacks and whites. Virtually none of the racial difference in earnings is

accounted for by differences in schooling, ability or experience levels. In terms of individual coefficients, the schooling estimates for blacks are slightly less than those for whites while the estimates of AFQT and experience effects are dramatically less for blacks.

Lest they be forgotten, however, a series of qualifications deserve a place at the end. The sample from which all of the analysis emanated is not a representative selection of the nation's population; instead it represents a group of draftees who left the Army during Fiscal Year 1969. Thus, the sample is fairly homogeneous with respect to age and experiences. It also contains less variation in schooling and ability levels than is found in the nation -- although there is still adequate representation of the tails of these distributions. The sample data for ten months after Army separation provides earnings levels at the very beginning of an individual's earnings profile. This necessitates some strong assumptions about the time path of earnings if generalizations about lifetime earnings functions are to be attempted. Finally, the sample refers to only individuals who are working full time, and, thus, the estimated relationships must be tempered by the probability that the individual is working.

Appendix A
Sample and Population Characteristics

Some feel for the extent of biases in sampling can be gained in Table A-1 from comparing sample characteristics with population characteristics on the distribution of years of schooling. The most notable difference between the sample and the population is that the sample distribution of education has smaller tails; there are proportionately neither as many individuals with college degrees nor with less than high school educations. The lower end is reduced by selectivity of the Army while the upper end is reduced by avoidance of the draft (e. g., with deferments), entry into the officer corps instead of the enlisted ranks, and the higher rate of entry into post-service educational programs by servicemen with higher education levels.[1/]

The Armed Force Qualification Test percentile distribution also gives some indication of the representativeness of the sample. The percentile scores are supposed to reflect the score distribution which would be obtained by the entire population. Table A-2 shows the means and tails of the sample distribution. The best comparison is, of course, between the white distribution and percentile norms since the norms

1/ Office of the Assistant Secretary of Defense (Manpower and Reserve Affairs), "Post-Service Educational, Training, and Vocational Status of First-Term Army Men Separating During July-September 1968," Manpower Research Reports, MA 70-1, October 1969.

Table A-1

Proportionate Years of School Completed:
Sample and Population

Race	Less than 12 years	12 years	13-15 years	16 or more years
Sample				
Total	.19	.60	.14	.07
White	.19	.59	.15	.07
Negro	.21	.65	.10	.04
National[a/]				
Total	.22	.43	.25	.10
White	.20	.43	.26	.11
Negro	.43	.38	.16	.03

a/ Source: Statistical Abstract of the United States, 1970, p. 111. The schooling figures for the nation apply to 1969 and are a weighted average of age 20-24 and age 25-29 statistics. The weighting (.8 and .2 respectively) reflects the relative age distribution in the sample.

Table A-2

AFQT Mean Percentiles and Distributions

Grouping	AFQT mean	Sample proporation AFQT percentile 0-30	93-100
Total	53.5	.266	.050
Urban	54.2	.273	.053
Rural	53.0	.260	.048
White	55.3	.225	.055
Urban	56.8	.209	.061
Rural	54.3	.235	.050
Black	28.7	.699	.001
Urban	30.4	.677	.001
Rural	26.3	.744	.000

most closely reflect white scores. The whites show a slightly higher than population mean, and some truncation in the two extremes.[1/] However, neither of these factors indicates a very serious distortion of the sample. The entire black distribution is shifted dramatically toward the lower end. There is, nevertheless, no good comparison to calculate whether or not this sample is representative of the population. Other people have noted that blacks tend to perform below national norms on these tests (Jensen[19], Karpinos[12]). The extent of this in the population is hard to judge.

Looking at the breakdown in Table A-3 of mean annual earnings within the sample also indicates that the sample is not out of line with the population as a whole. The median family income nationally for whites in 1968 was $8,937; for blacks, the comparable figure was $5,360.[2/] Given that all of the people in the sample are under 30 years of age and that these give incomes with ten months or less on the job, these sample values seem reasonable.

1/ The norms actually apply to a 1944 sample of the population; see Karpinos [12]. To the extent that AFQT scores reflect school quality and school quality has increased since 1944, the norms will be lower than actual distributions.

2/ U.S. Statistical Abstract, 1970, p. 324. The educational attainment of blacks in the sample (mean = 11.9) is considerably above the median for the country of 9.6 years. Median black income for a high school education was $6,432.

Table A-3

Sample Mean Annual Earnings[a/]

Grouping	
Total	$7,031
Urban	7,082
Rural	6,858
White	7,109
Urban	7,395
Rural	6,925
Black	6,199
Urban	6,458
Rural	5,677

a/ These calculations assume the individual is paid for 52 weeks.

REFERENCES

1. Becker, Gary S., Human Capital (New York: NBER, 1964)

2. Coleman, James S., et al, Equality of Educational Opportunity (Washington: GPO, 1966)

3. Fisher, Franklin M., "Tests of Equality Between Sets of Coefficients in Two Linear Regressions: An Expository Note", Econometrica (March, 1970)

4. Gintis, Herbert, "Education, Technology, and the Characteristics of Worker Productivity", American Economic Review (May, 1970)

5. Griliches, Zvi, "Notes on the Role of Education in Production Functions and Growth Accounting", in W. Lee Hansen (ed.) Education, Income, and Human Capital (New York: NBER, 1970)

6. _____ and W. Mason, "Education, Income and Ability", (mimeo, 1970)

7. Hanoch, Giora, "An Economic Analysis of Earnings and Schooling", Journal of Human Resources (Summer, 1967)

8. Hansen, W. Lee, Burton A. Weisbrod and William J. Scanlon, "Schooling and Earnings of Low Achievers", American Economic Review, (June 1970)

9. Hanushek, Eric A., "The Education of Negroes and Whites", (unpublished Ph. D. dissertation, M. I. T., 1968)

10. _____, "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data", American Economic Review (May, 1970)

11. _____, "The Regional Distribution of Returns to Human Capital", Program on Regional and Urban Economics Discussion Paper, Harvard University, 1971.

REFERENCES (con't)

12. Karpinos, Bernard D., "The Mental Qualifications of American Youths for Military Service and Its Relationship to Educational Attainment", 1966 Proceedings of the Social Statistics Section of the American Statistical Association, 1966.

13. Mincer, Jacob, "On-the-Job Training: Costs, Returns and Some Implications", Journal of Political Economy, Supplement, (October, 1962)

14. Samuelson, Paul A., "International Factor Price Equalization Once Again", The Economic Journal, (June, 1949)

15. Thurow, Lester C., Poverty and Discrimination (Washington: Brookings Institution, 1969)

16. Weiss, Randall D., "The Effects of Education on the Earnings of Blacks and Whites", Review of Economics and Statistics, (February, 1970)

17. Welch, Finis, "Measurement of the Quality of Schooling", American Economic Review (May, 1966)

18. Jensen, Arthur R., "How Much Can We Boost IQ and Scholastic Achievement?", Harvard Educational Review (Winter, 1969)